1Ш

УДК 004.89

А.И. Шевченко, О.А. Гудаев, С.П. Некрашевич
Государственный университет информатики и искусственного интеллекта,
г. Донецк, Украина

Проектирование системы мониторинга учебного процесса дистанционного образования на базе технологий искусственного интеллекта*

Описывается подход и используемые технологии для задач мониторинга производственного процесса дистанционного образования на основе онтологического подхода для представления данных информационного пространства. Характерными особенностями подхода являются шифрование данных маркерами расширенной реальности и реализация в виде мультиагентной системы.

Особенности предметной области

Рассмотрим документальный обмен информацией в среде мониторинга процессов дистанционного образования. Главный принцип функционирования схемы документооборота: информационное пространство строится в массовом некомпьютерном сообществе с разными уровнями освоения компьютерных технологий, с вырожденными коммуникационными связями, не поддающимися объединению технологиями компьютерных сетей. Государственные департаменты, студенты и преподаватели, администрация ВУЗа и муниципальные службы имеют разные уровни представления информации и сложности её обработки и в целом составляют массовое некомпьютерное сообщество, которое регулирует свою деятельность документальным способом. Цементирующим фактором обмена информацией в сообществе выступает печатный документ, который рассматривается через призму человеческого фактора, когда содержание может быть уточнено через новую копию. Такая система документооборота ориентирована на человека принимающего решение, который не спешит заполнять поля документа, относящиеся к рубрикации, версионированию и хронологии процесса обработки. Заполнение такого рода полей позволяет компьютеризировать построение информационного пространства. В действительности этого не происходит из-за непосильного объема рутинных работ, возлагаемых на автора документа. Взамен автору предлагается призрачная выгода от участия в цифровом обмене данными и автоматизированной обработке документов. В обществе не существует доверия к электронным способам передачи электронных данных, так как вероятность манипуляции с данными очень высока. Поэтому преобладает фиксация сведений о производственном процессе на материальных носителях. Многократно изменяя электронный документ, автор не заботится о контексте его создания. Место расположения электронного документа, версия и хронология модификаций являются контекстом создания. Документ в печатном

*Работа выполняется в рамках научной темы № 0109U002906 «Разработка интеллектуальной системы мониторинга учебного процесса дистанционного образования, которая использует шифрование управляющей информации маркерами расширенной реальности».

виде получает перечисленные характеристики естественным путём, приобретая пространственную и логическую локализацию, когда копии документов раскладываются по стопкам или папкам на столе.

Фундаментальный принцип организации документооборота

Учитывая особенности среды документооборота, можно предложить следующий подход организации информационного пространства с применением технологий искусственного интеллекта [1].

Реальная конфигурация компьютерного класса, домашнего компьютера студента и рабочего места преподавателя не предполагают постоянного подключения к системе. Поэтому выполнение и сбор учебных заданий по дисциплине разнесен в пространстве и времени. Рассмотрим три варианта организации учебного процесса на базе компьютерных технологий.

Вариант «А». Студент выполняет задание в компьютерном классе, используя терминал системы. Недостатки: отсутствует работа с преподавателем над примерами, необходимость мониторинга лектором результатов работы с применением компьютера или затраты на печать результатов ассистентом, необходимость переноса электронного задания из системы на домашний компьютер студента.

Вариант «Б». Студент получает задание в виде исполняемого модуля и копирует его в классе на запоминающее устройство Flash-памяти. Студент выполняет задание в часы самоподготовки в домашних условиях, в условиях дополнительной работы в классе или в библиотеке. Возможна консультация с преподавателем и работа над ошибками. Сведения о результатах работы над заданием студент сдает ассистенту в электронном виде. Недостатки: сложность в организации сбора электронных данных ассистентом и потеря времени на практических занятиях на это, затраты на оперативность анализа лектором результатов, как в варианте «А».

Вариант «С». Студент получает задание как в варианте «Б», но результат печатает в виде специальной формы документа с электронной подписью. Ассистент на практическом занятии оперативно собирает документальные сведения, затрачивая минимум организационных усилий, так как документ не требует регистрации и проверки его корректности принимающим лицом. В случае непредвиденных ситуаций (медосмотр, каникулы, перенос занятий, тревога гражданской обороны, болезни студента) документ может быт передан (старостой группы, деканом, родителями) лектору любым способом (почтой или курьером). Лектор оперативно и без дополнительных средств, рассмотрев печатные документы, может оценить объем выполненных работ ассистентом и студентами. Для анализа результатов лектор организует ввод документа в систему в виде растрового изображения по средствам планшетного сканера или мобильного телефона, поручая эту работу любому сотруднику учреждения, умеющему работать с копировальной техникой.

Многолетняя практика компьютерных комплексов ввода печатных документов показала, что с функциями сканирования может справиться любой низкоквалифицированный работник без специального компьютерного образования и навыков работы, если он умеет переключать пультом дистанционного управления программы на телевизоре.

Организовав цепочку технологического процесса сбора данных, разнесенного в пространстве и времени, лектор добивается тотального контроля над выполнением операций и полного понимания участниками процесса его организации.

Недостатки: небольшой объем учебных заданий, затруднена работа студентов по исправлению ошибок в результатах и затраты на печать нового ответа, затраты на получения графических изображений документов.

Использование варианта «С» позволяет увеличить количество проверяемых преподавателем заданий за счет автоматизации обработки информации и формирование первичных данных, вводимых в систему, на месте их образования. Поэтому недостаток небольшого объема учебного задания нивелируется, если увеличить количество заданий.

Небольшой объем задания служит предпосылкой к повышению качества образования: преподаватель не может разработать сложные путанные задания и студент работает с небольшими порциями материала. Небольшая порция материала в задании порождает меньшее количество ошибок. В случае если ошибки присутствуют, то они несущественно влияют на общую картину мониторинга образовательного процесса, так как аналогичное задание не сложно выполнить повторно.

Затраты на неквалифицированный труд по вводу документов могут быть распределены между лектором, ассистентом, методистом и практикантами, позволив сконцентрироваться на разработке серии заданий.

Вывод. Ценности Болонского процесса разработаны высокоразвитыми странами, как эволюция корпоративных АСУ университетов, когда компьютерная обработка данных учебного процесса автоматизирована полностью. Построенная по варианту «С» распределенная система искусственного интеллекта учитывает возможности существующей инфраструктуры образования в Украине и только тогда раскрывает в полном объеме ценности Болонского процесса.

Проектирование архитектуры

Система проектируется как архитектура Клиент/Сервер (рис. 1). Клиентами являются удаленные пользователи системы с непостоянным подключением. Клиент получает обновления некоторыми порциями данных и размещает в локальном хранилище. Принцип организации хранилища позволяет репликацию данных самым простым методом: порция данных дописывается в файл как новый слой семантической сети.

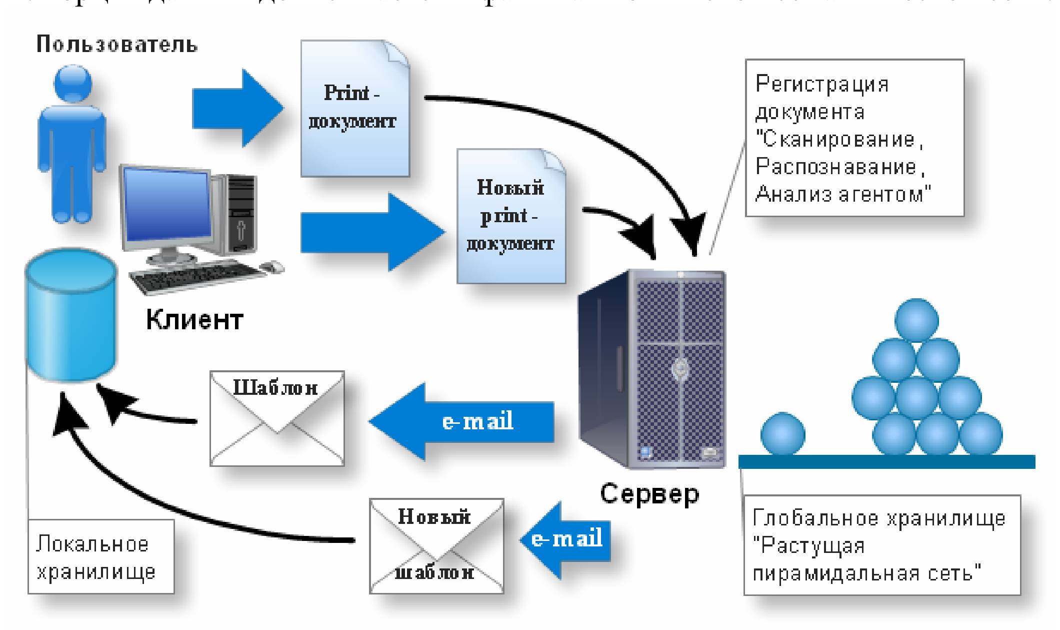


Рисунок 1 – Архитектура системы

В системе сервер выступает в качестве центра распределения цифровых подписей. Центр распределения накапливает сведения об учебном процессе в онтологическом хранилище. Центр генерирует шаблон документа и модель кодирования электронных сведений, которые размещаются на печатном документе [2]. Для удалённой работы с локальным хранилищем клиента сервер организует пакетную передачу новых сведений об учебном процессе. Новые данные могут представлять собой шаблоны документов с указанием словаря кодирования маркерами электронной подписи [3].

Сбор и обработка данных, циркулирующих в системе, имеет уникальную схему. Уникальность схемы заключается в использовании комбинации методов искусственного интеллекта.

В начале работы системы глобальное хранилище сервера имеет первоначальный шаблон документа, который предполагается заполнять клиенту. Шаблон документа представляется семантической сетью и хранится в виде пакета синхронизации [4], [5]. Глобальное хранилище не содержит сведений о клиентах и имеет самое общее описание структуры предметной области без заполнения базы данных. База данных на нижнем уровне представления имеет сетевую модель хранения данных, а на предметном уровне имеет смешанную модель хранения гетерогенных сведений. В предметной базе данных можно хранить сеть онтологий и реляционные таблицы с данными. Например, такой подход развития нереляционных способов ведения глобальных хранилищ преобладает в современной системе Facebook [6].

Пользователь программного модуля клиента заполняет бланк документа. Например, документ содержит поля «Фамилия» и «Учебная группа». К документу прикрепляется электронная подпись, содержащая заполненные поля. Клиент электронную подпись кодирует графическим способом с помощью маркеров расширенной реальности [2], [3]. Правила кодирования являются неотъемлемой частью шаблона документа. Правила описываются семантической сетью онтологий. Графическое изображение кода размещается в некоторой зоне печатного документа. Кодирование электронных данных маркерами и размещение их на печатных документах позволяет автоматизировать обработку документа сервером. Электронная подпись на печатном документе решает следующие задачи: регистрация документа, разрешение конфликтов дублирования документов, составление временной хронологии и версии документа. Самое главное, электронная подпись позволяет автоматизировать ввод первичных данных в автоматизированную систему. Таким образом используется децентрализованная модель сбора данных. Пользователь клиента выполняет роль оператора системы. Где зарождается информация, там и организован её ввод в систему. Печатный документ выполняет роль транспортного носителя электронных данных при передаче системе.

Печатный документ содержит поля, которые заполнены вручную. Например, рецензию руководителя или ответы на вопросы теста. В рассматриваемой системе графическое изображение печатного документа попадает в графическое хранилище сервера. Используя методы распознавания графических образов, электронная подпись документа расшифровывается [3]. Извлеченные из документа электронные данные обрабатываются интеллектуальным агентом и в виде семантической сети сохраняются в глобальное хранилище сервера. Теперь система больше знает о пользователе. Система использует полученные сведения о пользователе для мониторинга учебного процесса.

Система работает по принципу растущей пирамидальной сети, когда пополняет семантическую сеть глобального хранилища сведениями о конечном пользователе. Следующий шаг в обработке данных делают эксперт предметной области и инженер по знаниям. Они могут составить новый шаблон документа на основании рекомендаций интеллектуального агента системы. Например, агент, сделав логический вывод в онтологическом пространстве хранилища, предлагает сформировать журнал студентов,

объединив их по признаку принадлежности к учебной группе. Журнал – это реляционное представление списка студентов, каждая запись которого может быть закодирована первичным ключом. Ключи остаются узлами семантической сети. Теперь новый шаблон цифровой подписи документа вместо поля «Фамилия» содержит значение ключа в журнале студентов. Тогда возникает задача: как передать пользователю новый шаблон документа и его персональный ключ, сгенерированный центром распределения цифровых подписей. Сервер формирует пакет обновления локального хранилища пользователя, которое легко передать по сети Интернет.

Такой семантический приём кодирования уменьшает объем кодируемого сообщения. Это важно в условиях дефицита свободного пространства для размещения цифровой подписи на печатном документе. Но самое главное, семантический приём нормализует информацию о пользователе для принятия решений агентами системы и для использования OLAP-анализа. Таким образом, шаблон цифровой подписи документа изменяется по мере наполнения онтологического хранилища.

Вывод. Управляя эволюционным ростом онтологического хранилища, система планирует акции взаимодействия с пользователем. У пользователя создается впечатление об персонификации интерфейса. Уточняя сведения о пользователе, система эволюционирует на основании объективных условий учебного процесса: отсутствует отвлеченность от предметной области проектировщика, а в случае его заблуждения – отвлеченность отмирает на новом этапе обновления шаблонов документа, как неиспользуемый балласт.

Документно-ориентированный подход

Предметная область (ПрО) дистанционного образования содержит понятия, центральное место среди которых занимают документы различных типов. В отличие от классической системы документооборота, эти документы являются открытыми, распределенными во времени и пространстве, федеративными (интегрируют данные, ресурсы и сервисы из различных источников), противоречивыми (обладают неполнотой и многозначностью). Характерной особенностью документов является то, что они содержат описательную часть, понятную как человеку, так и компьютеру.

Документ содержит часть информации, закодированной с использованием маркеров расширенной реальности. Таким образом, информация, представленная на естественном языке, с использованием маркеров расширенной реальности представляется в удобном для машинной обработки формате, а с использованием специальных аудио и визуализаторов, доступна и для человека.

Автором каждого документа является человек-пользователь системы, который внес первичную информацию (данные) в документ. Соавтором документа является человек, который добавляет дополнительную информацию, например результаты оценок или аттестации.

Каждый документ имеет определенную структуру (схему), задаваемую разработчиком системы. Тогда будем говорить об авторе схемы документа, им обычно является разработчик.

Современной тенденцией в разработке распределенных информационных систем является уменьшение времени разработки систем. Часто пользователи вовлекаются в проектирование и совместное построение системы. Таким образом, значительно сокращается время до введения системы в эксплуатацию, а также становится возможным функционирование системы еще на этапе прототипирования. Важной задачей при таких подходах является поддержка целостности данных при версионном изменении системы.

Формализация производственного процесса обработки документов

Информация, содержащаяся в документах, имеет смысл в терминах рабочего процесса, она определяет действия, которые необходимо предпринять по обработке и дальнейшему использованию в рабочем процессе.

Таким образом, работа с документами в системе сводится к управлению задачами. Появление в системе нового или обновленного документа приводит к генерации новых задач.

Для каждого типа документов вводится свой жизненный цикл задач, который можно формально представить в виде UML диаграмм состояний (рис. 2).

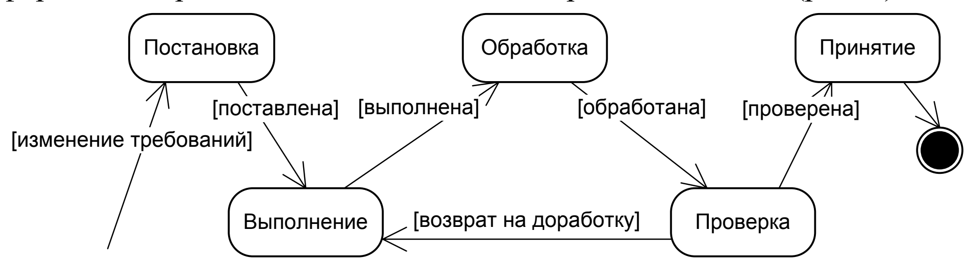


Рисунок 2 – Жизненный цикл документа

В любой момент времени документ может находиться только в одном из множества состояний. С определенной вероятностью документ может перейти на другой этап обработки.

Каждая задача является элементарным действием. Для выполнения задачи назначается 1 или более исполнителей. При количестве исполнителей более 1 необходимо организовать согласования для совместного выполнения задачи.

Для проверки задачи задается один или несколько контроллеров. При количестве контроллеров более 1 необходимо организовать согласования для совместного принятия задачи. Согласования при проверке и выполнении отличаются видами деятельности, и часто первое может выполняться параллельно, в то время как выполнение может требовать определенной последовательности действий.

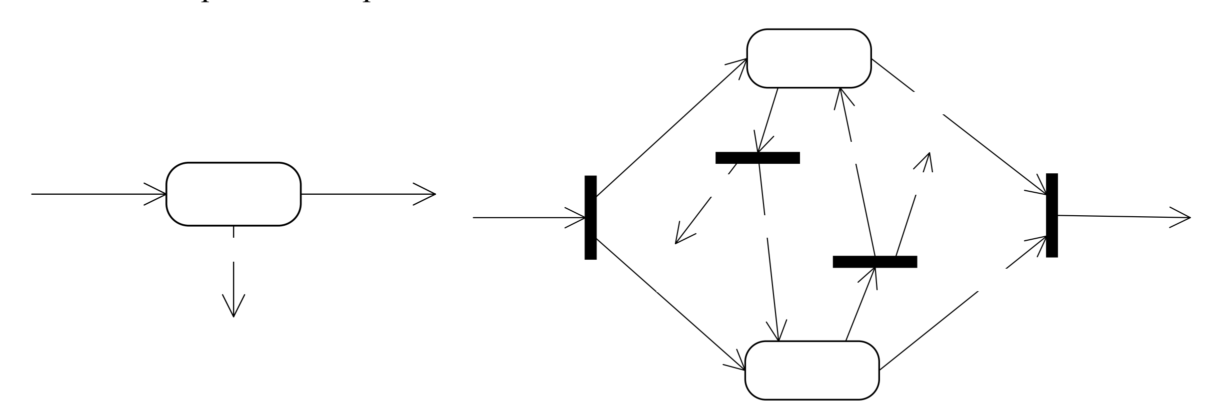


Рисунок 3 – Выполнение этапа одним и двумя исполнителями

Если отдельный этап производственного процесса является композитным, например, состоит из нескольких независимых подзадач или требует вовлечения нескольких исполнителей, то диаграмму состояний документа можно модифицировать по схеме, приведенной на рис. 3, включающей разветвления работ и блокировки параллельных процессов для ожидания их совместного выполнения.

Реализация мультиагентной системы

Для реализации предлагается использовать мультиагентную систему (МАС), состоящую из агентов различных видов.

1. Агент анализа и обработки документа.
2. Агент взаимодействия и планирования.
3. Агент экспертной аттестации документа.
4. Агенты мониторинга параметров ПрО.
5. Агент коммуникаций.

Агенты являются программными модулями, которые предоставляют определенные сервисы [7], [8]. Агенты используют данные и знания, представленные в хранилищах. В качестве основного интерфейса доступа к данным используется декларативный язык запросов SPARQL. Также агенты обладают собственными данными и знаниями и действуют по ситуации, то есть являются нерефлексивными.

Агент анализа и обработки документа выполняет диспетчеризацию обработки первичного документа, позволяет осуществить связь между физическими документами различных типов и версий и семантическим представлением информации, которая используется в задачах мониторинга предметной области:

1. Получение документа (1 лист, несколько листов, несколько документов).
2. Обнаружение маркеров.
3. Получение семантики документа (анализ маркеров).
4. Аудио и визуализация результатов.
5. Предварительная обработка документа – расширенный анализ документа с использованием агентов.
6. Постобработка документа – событийный запуск агентов с возможным человеко-машинным интерфейсом.

Агент взаимодействия и планирования выполняет координацию работы агентов, а также взаимодействует с человеком как специалистом в разработке, так и пользователем системы. Агент выполняет следующие задачи:

1. Формирование задач на обработку и аттестацию документов.
2. Подготовка новых версий документов.
3. Оценивание приоритета задач, формирование очереди для одного исполнителя.
4. Планирование встреч, занятий, экзаменов.

Агент экспертной аттестации документа выполняет проверку и принятие документа, выводит и добавляет новые знания в онтологию (оценки, рекомендации, стратегии), выполняет следующие задачи:

1. Поиск аналогичных документов (подобие, заимствование).
2. Классификация и кластеризация документов.
3. Определение структуры документа.
4. Извлечение программной части.
5. Оценивание работы.
6. Валидация работы (обнаружение ошибок, несоответствий, отклонений).
7. Получение метрик документа (объем, количество страниц, графики, ссылки).

Агенты мониторинга параметров ПрО выполняют интеллектуальный анализ данных в режиме онлайн на основе онтологии, содержащей представление семантики данных о документах в системе. Агент предоставляет следующие сервисы:

1. Интеллектуальный онтолого-ориентированный поиск данных.
2. OLAP-анализ.

3. Датамайнинг – выявление скрытых трендов.
4. Кластеризация.
5. Прогнозирование.
6. Планирование и выбор стратегий.

Агент мониторинга параметров ПрО использует текущее состояние документов в производственном процессе, оперируя следующими метриками:

1. % документов, находящихся в определенном состоянии (обрабатываются, на проверке).
2. Вероятности возврата документа с определенного этапа.
3. Профиль документа и совокупности для определенного автора.
4. Профиль загрузки исполнителей.

Агент коммуникаций используется для гарантированной доставки документа адресату. Агент является независимым от транспортного протокола: в каждом конкретном случае может использоваться электронная почта, веб-интерфейс, СМС и другие коммуникационные способы передачи сообщений. Поскольку постоянное соединение клиентов с агентом отсутствует, то необходимо использовать промежуточное ПО очереди сообщений для гарантированной доставки сообщения.

Кодирование маркерами данных документа

Когда система начинает работать с пользователем, то предоставляет ему типовой шаблон документа. Документ в клиенте представлен графической формой ввода данных. Заполненные поля формы будут отражены в электронной подписи маркерами.

Широко известные форматы файлов PDF и XML используют произвольный порядок полей. В них для обозначения структурных элементов документа используют метки. Метка является идентификатором значения в потоке данных. В языке представления документов XML в качестве метки используют символьный тег. В бинарных спецификациях для представления меток используют числовой код. В настольных базах данных обходятся без меток, так как знают размер всех полей записи.

В электронной подписи документа метку можно представить одним набором маркеров, а значения – другим набором маркеров (рис. 4). Используя несколько кодовых алфавитов маркеров (табл. 1), можно уменьшит длину электронной подписи. Одни маркеры обозначают номер поля, другие – 8 бит данных, а третьи – символы текста в ASCII-кодировке.

Из представленной на рис. 4 семантической сети документа можно составить шаблон электронной подписи, имеющий следующее формальное представление:

```
01,02(4e402fb3de7e447ea68144cd92f11a05),
03(662ed1b21b484367a79a96a78bf34508),
04(4440f57ae54c4494b8b5133d3b68a7b6),
05("Бобров"), 06("Кирилл"), 07("Олегович"), 08("ПО07а"),
09("2009").
```

Таблица 1 – Распределение маркеров

Ко-во	Диапазон маркеров	Назначение
50	[1..50]	Обозначение полей семантической сети
256	[51-307]	Кодирование цифр 256-ричной системы счисления
256	[308-564]	Кодирование ASCII-символов

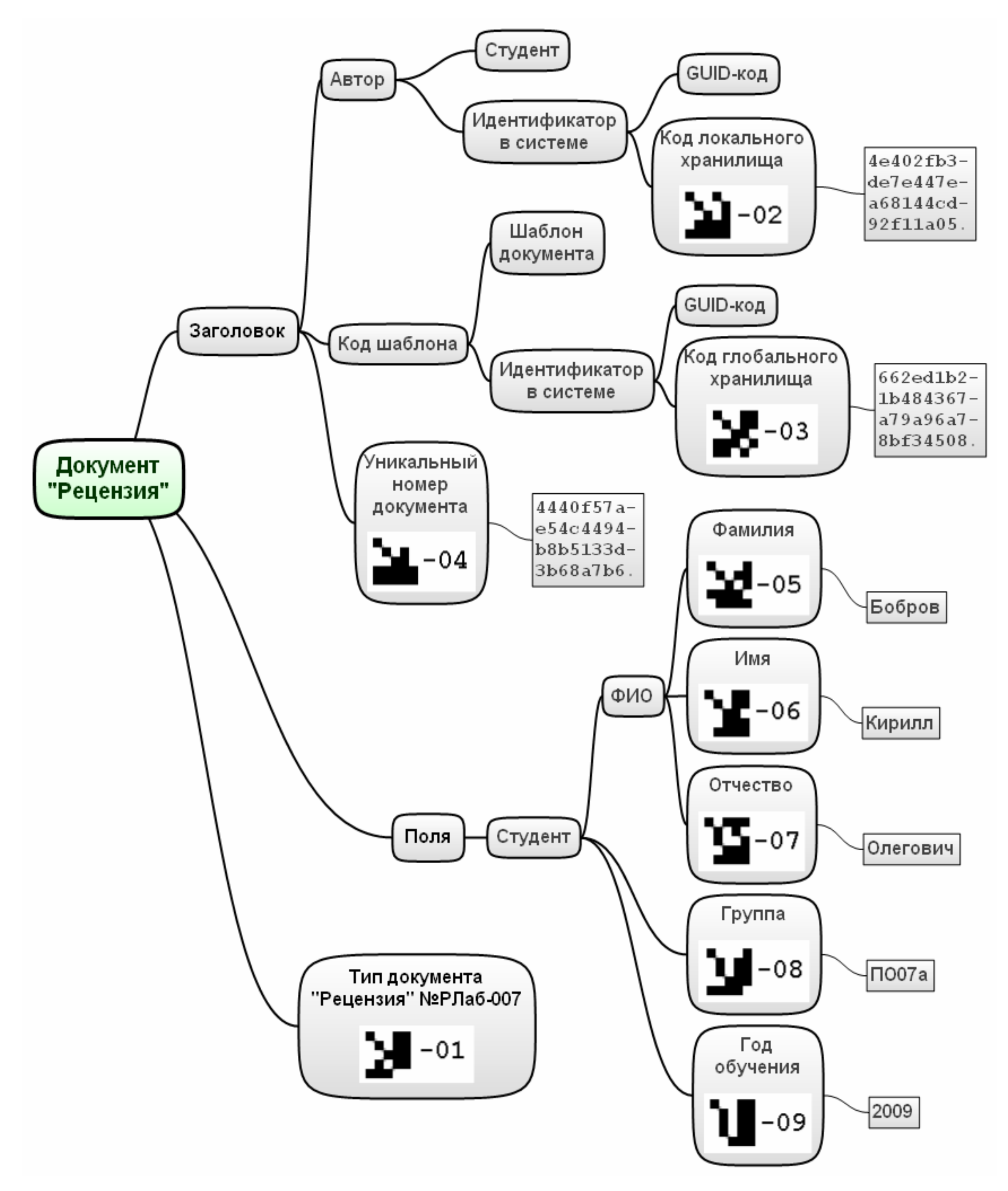


Рисунок 4 – Кодирование семантической сети маркерами

Выводы

В работе описывается структура и бизнес-процессы системы мониторинга учебного процесса дистанционного образования. Разработана функциональная модель на основе мультиагентного подхода. Предложен формат представления семантики документа в виде семантической сети. Для представления информации в документе, подлежащей компьютерной обработке мультиагентной системой, используются маркеры расширенной реальности. Посредством использования предложенных подходов

решается вопрос интеграции сервисов предметной области. Таким образом, становится возможным переход от автоматизации рабочих процессов дистанционного образования к интеллектуальной обработке с использованием аппарата искусственного интеллекта.

Литература

1. Шевченко А.И. Актуальные проблемы теории искусственного интеллекта: [Монографія] / Шевченко А.И. – Київ : ІПШІ «Наука і освіта», 2003. – 226 с.
2. Гудаев О.А. Аффинные углы поворота плоскости маркера расширенной реальности ARGET, используемые в интеллектуальной обработке информации системой дистанционного образования / О.А. Гудаев, А.И. Шевченко // IX Международная научная конференция имени Т.А. Таран «Интеллектуальный анализ информации ИАИ-2009» (Киев, 19 – 22 мая 2009 г.): сб. тр. / [ред. кол. : С.В. Сирота (гл. ред.) и др.] – К. : Просвіта, 2009. – С. 87-93.
3. Гудаев О.А. Синтез и анализ предложений графического языка передачи сообщений в мобильных робототехнических системах с элементами расширенной реальности (ARGET) / О.А. Гудаев // Искусственный интеллект. – 2006. – № 2. – С. 467-498.
4. Некрашевич С.П. Представление данных в Интернет на основе семантических сетей / С.П. Некрашевич, Д.В. Божко // Искусственный интеллект. – 2006. – № 1. – С. 57-65.
5. Некрашевич С.П. Построение модели онтологии интеллектуальной системы мониторинга учебного процесса дистанционного образования / С.П. Некрашевич // Искусственный интеллект. – 2009. – № 2. – С. 124-129.
6. Зубинский А. Время других DBMS / А. Зубинский // Компьютерное Обозрение. – 19 – 25 мая 2009. – № 16 – 17 (683). – С. 48-50.
7. Некрашевич С.П. Агентно-ориентированный подход к разработке сложных программных систем / С.П. Некрашевич // Вестник НТУ«ХПИ». – 2004. – № 36.
8. Некрашевич С.П. Интеллектуализация средств предоставления бизнес-услуг на основе сети Интернет / С.П. Некрашевич, И.Н. Рубцова // Искусственный интеллект. – 2006. – № 1. – С. 66-70.

А.І. Шевченко, О.О. Гудаєв, С.П. Некрашевич
Проектування системи моніторингу навчального процесу дистанційної освіти на базі технологій штучного інтелекту
Описується підхід і використання технології для моніторингу виробничого процесу дистанційної освіти на основі онтологічного підходу для представлення даних інформаційного простору. Важливими особливостями підходу є шифрування даних маркерами розширеної реальності і реалізація у вигляді мультіагентної системи.

A.I. Shevchenko, O.O. Gudayev, S.P. Necrashevich
Design of Monitoring System of Distance Education Teaching on the Basis of Artificial Intelligence Technologies
The approach and in-use technologies are described for the tasks of monitoring of production process distance education on the basis of ontological approach for presentation of information. The characteristic features of the approach are coding of information and realization the markers of the Augmented Reality as a multi-agent.

Статья поступила в редакцию 25.01.2010.